C. ASSESSMENT OF GOLDEN TILEFISH (*Lopholatilus chamaeleonticeps*) in the Middle Atlantic-Southern New England Region

A Report of the
Southern Demersal Working Group
National Marine Fisheries Service
Northeast Fisheries Science Center
Woods Hole, MA 02543

EXECUTIVE SUMMARY

Terms of Reference (TOR):

1. *Characterize the commercial catch including landings and discards. Characterize recreational landings.*
This TOR was completed. See Section 2.0.

2. *Estimate fishing mortality and total stock biomass for the current year and characterize the uncertainty of those estimates*.
This TOR was completed. See Section 3.0.

3. *Evaluate and either update or re-estimate biological reference points as appropriate*.
This TOR was completed. See Section 3.0.

4. *Where appropriate, estimate a constant TAC and/or TAL based on stock status for years following the terminal assessment year.* This TOR is covered in TOR 5.

5. *If projections are possible,*

a) *provide seven year projections of stock status under various TAC strategies and*

b) *evaluate current and projected stock status against existing rebuilding or recovery schedules, as appropriate.*

This TOR was not carried out because of concerns related to the wide variance and substantial bias in the projection realizations. See Section 4.0.

6. *Review, evaluate and report on the status of the research recommendations offered in the 1999 Science and Statistical committee reviewed assessment.*
This TOR was completed. See Section 7.0.

The current status for this stock is based on the ASPIC surplus production model employed in the past 2 assessments. The model is calibrated with CPUE series, as there are no fishery-independent sources of information on trends in population abundance. While the Working Group expressed concern about the projection phase of this analysis, we agreed to accept the estimates of current fishing mortality and biomass and associated reference points.

Total commercial landings (live weight) increased from less than 125 metric tons (mt) during 1967-1972 to more than 3,900 mt in 1979 and 1980. Annual landings have ranged between 666 and 1,838 mt from 1988 to 1998. Landings from 1999 to 2002 were below 900 mt (ranging from 506 to 874 mt). An annual quota of 905 mt was implemented in November of 2001. During the late 1970s and early 1980s Barnegat, NJ was the principal tilefish port; more recently Montauk, NY has accounted for most of the landings.

Three different series of longline effort data were analyzed. The first series was developed by Turner (1986) who used a general linear modeling approach to standardize tilefish effort during 1973-1982 measured in kg per tub (0.9 km of groundline with a hook every 3.7 m) of longline fished obtained from logbooks of tilefish fishermen. Two additional CPUE series were calculated from the NEFSC weighout (1979-1993) and the VTR (1995-2004) systems. The number of vessels targeting tilefish has declined over the time series; during 1995-2002, five vessels accounted for more than 70 percent of the total tilefish landings. The length of a targeted tilefish trip had been generally increasing until the mid 1990s. Since then there appears to have been a trend towards shortening of the tilefish trips.

Six market categories exist in the database. From smallest to largest they are: small, kitten, medium, large and extra large as well as an unclassified category. The proportion of landings in the kittens and small market categories increased in 1995 and 1996. Evidence of two strong recruitment events can be seen tracking through these market categories. The proportion of large market category has declined since the early 1980s. Commercial length sampling has been inadequate over most of the time series. However some commercial length sampling occurred in the mid to late 1990s. More recently there has been a substantial increase in the commercial length sampling in 2003 and 2004.

A small recreational fishery occurred briefly in the mid 1970s (< 100 mt annually) but subsequent recreational catches have been quite low for the last 25 years (i.e., less than 1 mt caught annually). Directed tilefish trips are rare. Since 2000, only 2 trips in the MRFSS data had tilefish reported as the primary target species.

Thirteen different configurations of the ASPIC model were examined. The accepted formulation began the analysis in 1973, separated the Turner, weighout and VTR CPUE into three series and fixed the $B1/B_{msy}$ ratio at 1 as the final run (run 13). The surplus production model indicates that the tilefish stock biomass in 2005 has improved since the last assessment in 1998. Total biomass in 2005 is estimated to be 72% of B_{msy} and fishing mortality in 2004 is estimated to be 87% of F_{msy}. Biological reference points did not change greatly from the 1998 assessment. B_{msy} is estimated to be 9,384 mt and F_{msy} is estimated to be 0.21.

Results from several alternative models were also examined. Results from An Index Method (AIM) model also suggest that relative F is below the point that corresponds with a replacement ratio of 1 (stock replacement). MSY and Yield per recruit based biological reference points did not change greatly from the 1998 assessment. The Lagged Recruitment Survival Growth (LRSG) model produced results similar to the ASPIC surplus production model calibrated with the single linked CPUE series. However commercial length data indicate that increases in total biomass are predominantly due to a strong 1999 year class. Most of the commercial catch over the 2002-2004 period was derived from this year class.

Several ASPIC projections employing a constant TAC strategy, including the current TAC of 905 mt were examined Each of these analyses exhibited wide variance and substantial bias and, in many cases, produced estimates of biomass and F at maximum or minimum model boundary conditions. The projections are too uncertain to form the basis for evaluating likely biomass recovery schedules relative to B_{msy} under various TAC strategies. The Working Group does note, however, that stock biomass in 2005 (72% of B_{msy}) is above that projected for 2005 in the 1998 assessment (59% of B_{msy}). Thus, the existing TAC of 905 mt appears to have sufficiently constrained F to allow stock biomass to increase towards B_{msy}.

There are two major sources of uncertainty affecting our perception of current stock status. The biomass-based models (ASPIC, AIM and LRSG) use the CPUE series as an index of population size. The Working Group considered these models and expressed concerns over whether the CPUE in this fishery may be as much a reflection of changes in fishing practices and changes in spatial distribution of the fish rather than fluctuations in population size. The catch-length model attempts to reconcile recent fishing mortality rates with a less than expected representation of larger fish in the catch. Because there are no fishery-independent data on trends in population biomass and size structure, the model must assume that the length composition of the catch will represent the extent of large fish in the population assuming a flat topped partial recruitment pattern. Working Group comments are included as Appendix C1.

1.0 INTRODUCTION

Golden tilefish, *Lopholatilus chamaeleonticeps*, inhabit the outer continental shelf from Nova Scotia to South America, and are relatively abundant in the Southern New England to Mid-Atlantic region at depths of 80 to 440 m. Tilefish have a narrow temperature preference of 9 to 14 C. Their temperature preference limits their range to a narrow band along the upper slope of the continental shelf where temperatures vary by only a few degrees over the year. They are generally found in and around submarine canyons where they occupy burrows in the sedimentary substrate. Tilefish are relatively slow growing and long-lived, with a maximum observed age of 46 years and a maximum length of 110 cm for females and 39 years and 112 cm for males (Turner 1986). At lengths exceeding 70 cm, the predorsal adipose flap, characteristic of this species, is larger in males and can be used to distinguish the sexes. Tilefish of both sexes are mature at ages between 5 and 7 years (Grimes et. al. 1988).

Golden Tilefish was first assessed at SARC 16 in 1992 (NEFSC 1993). The Stock Assessment Review Committee (SARC) accepted a non-equilibrium surplus production model (ASPIC). The ASPIC model estimated biomass-based fishing mortality (F) in 1992 to be 3-times higher than F_{msy}, and the 1992 total stock biomass to be about 40% of B_{msy}. The intrinsic rate of increase (r) was estimated at 0.22.

The Science and Statistical (S&S) Committee reviewed an updated tilefish assessment in 1999. Total biomass in 1998 was estimated to be 2,936 mt, which was 35% of B_{msy} = 8,448 mt. Fishing mortality was estimated to be 0.45 in 1998, which was about 2-times higher than F_{msy} = 0.22. The intrinsic rate of increase (r) was estimated to be 0.45. These results were used in the development of the Tilefish Fishery Management Plan (Mid-Atlantic Fishery Management Council 2000). The Mid-Atlantic Fishery Management Council implemented the Tilefish Fishery Management Plan (FMP) in November of 2001. Rebuilding of the tilefish stock to B_{msy} was based on a ten-year constant harvest quota of 905 mt.

TOR 1*: Characterize the commercial catch including landings and discards. Characterize recreational landings.*

2.0 DATA SOURCES

Commercial catch data

Total commercial landings (live weight) increased from less than 125 mt during 1967-1972 to more than 3,900 mt in 1979 and 1980 (Table C1, Figure C1). Landings stabilized at about 2,000 mt during 1982-1986. An increase in landings occurred in 1987 to 3,200 mt but subsequently declined to 450 mt in 1989. Annual landings have ranged between 454 and 1,838 mt from 1988 to 1998. Landings from 1999 to 2002 were below 900 mt (ranging from 506 to 874 mt). An annual quota of 905 mt was implemented in November of 2001. Landings in 2003 and 2004 were over the quota at 1,130 and 1,182 mt respectively. Over 75% of the landings came from Statistical Areas 537 and 616 since

1991 (Table C2). Since the 1980s, over 85% of the commercial landings of tilefish in the MA-SNE region have been taken in the longline fishery (Table C3, Figure C2). During the late 1970s and early 1980s Barnegat, NJ was the principal tilefish port; more recently Montauk, NY has accounted for most of the landings. The shift in landings can be seen in the proportion of the landings by state in Table C4 and Figure C3. In the late 1970s and earlier 1980s a greater proportion of the landings were taken in quarters 1 and 2 (Table C5, Figure C4). Recent landings have been relatively constant over the year.

Commercial discard data

Very little discarding (< 1%) of tilefish was reported in the vessel trip report (VTR) from longline vessels that target tilefish and there is little reported discarding of tilefish in the trawl fishery in the VTR data (Table C6). The highest trawl reported total discard of tilefish was 13 mt in 2003. Observer trawl data did not produce a reliable discard estimates for tilefish. Discard to kept ratios for trawl trips that either kept or discarded tilefish in the observer data varied from 0 in 1993 to 1.4 in 2001 (Table C7). Since 1989, twelve of the sixteen years had less than 15 trips sampled that caught tilefish.

Commercial CPUE data

Analyses of catch (landings) and effort data were confined to the longline fishery since directed tilefish effort occurs in this fishery (e.g. the remainder of tilefish landings are taken as bycatch in the trawl fishery). Most longline trips that catch tilefish fall into two categories: (a) trips in which tilefish comprise greater than 90% of the trip catch by weight and (b) trips in which tilefish accounted for less than 10% of the catch. Effort was considered directed for tilefish when at least 75% of the catch from a trip consisted of tilefish (NEFSC 1993).

Three different series of longline effort data were analyzed. The first series was developed by Turner (1986) who used a general linear modeling approach to standardize tilefish effort during 1973-1982 measured in kg per tub (0.9 km of groundline with a hook every 3.7 m) of longline obtained from logbooks of tilefish fishermen. Two additional CPUE series were calculated from the NEFSC weighout (1979-1993) and the VTR (1995-2004) systems as well as a combined 1979-2004 series. Effort from the weighout data was derived by port agents' interviews with vessel captains whereas effort from the VTR systems comes directly from mandatory logbook data. In this assessment and in the 1998 tilefish assessment we used Days absent as the best available effort metric. In the 1998 assessment an effort metric based on Days fished (average hours fished per set / 24 * number of sets in trip) was not used because effort data were missing in many of the logbooks and the effort data were collected on a trip basis as opposed to a haul by haul basis. For this assessment effort was calculated as:

$$\text{Effort} = \text{Days absent} - \text{Number of trips},$$

where, Days absent = (time & date landed - time & date sailed).
For some trips, the reported days absent were calculated to be a single day. This was considered unlikely, as a directed tilefish trip requires time for a vessel to steam to near the edge of the continental shelf, time for fishing, and return trip time (Grimes et al.

1980). Thus, to produce a realistic effort metric based on days absent, a one day steam time for each trip (or the number of trips) was subtracted from days absents and therefore only trips with days absent greater than one day were used.

The NEFSC Weighout and VTR CPUE series were standardized using a general linear model (GLM) incorporating year and individual vessel effects (Mayo et al. 1994). The CPUE was standardized to an individual longline vessel and the year 1984; the same year used in the last assessment. For the VTR series the year 2000 was used as the standard. Model coefficients were back-transformed to a linear scale after correcting for transformation bias (Granger and Newbold 1977). The full GLM output for the Weighout CPUE series is included as Appendix C2 and the full GLM output for the VTR CPUE series is included as Appendix C3.

The number of vessels targeting tilefish has declined over the time series (Table C8, Figure C5); during 1995-2002, five vessels accounted for more than 70 percent of the total tilefish landings (Table C9, Figure C6). In 2003 and 2004 there appears to be an increase in the number of vessels targeting tilefish. The length of a targeted tilefish trip had been generally increasing until the mid 1990s. Since then there appears to have been a trend towards decreasing trip length (Figure C5). In the weighout data the small number of interview is a source of concern; very little interview data exists at the beginning of the time series (Table C8, Figure C7). The 5 dominant tilefish vessels make up almost all of the VTR data with the exception of 2004 when there appears to be more vessels targeting tilefish (Figure C6). In some years there were higher total landings reported in the VTR data than the Dealer data for the 5 dominant tilefish vessels.

The number of targeted tilefish trips declined in the early 1980s while trip length increased (Figures C5 and C8). More recently the number of trips became relatively stable as trip length decreased. The interaction between the number of vessels, the length of a trip and the number of trips can be seen in the total days absent trend in Figure C8. Total days absent remained relatively stable in the early 1980s, but then declined at the end of the weighout series (1979-1994). In the beginning of the VTR series (1994-2004) days absent increased through 1998 but declined thereafter. Figure C8 also shows that a smaller fraction of the total landings were included in the calculation of CPUE compared to the VTR series.

Figure C9 illustrates difference between the nominal CPUE and vessel standardized (GLM) CPUE with the weighout and VTR data combined. A large increase in CPUE can be seen in both series in recent years. CPUE trends are similar for most vessels that targeted tilefish (Figure C10). The sensitivity of the GLM model to sporadic vessels entering the CPUE series was tested by limiting the CPUE data set to vessels that were represented for at least 2 years, 3 years, 4 years, 5 years, and 6 years (Figures C11 to C15). This trimming of the data had very little influence on the resulting standardized GLM CPUE trend (Figure C16).

Very little CPUE data exist for New York vessels in the 1979-1994 weighout series despite the shift in landing from New Jersey to New York before the start of the VTR series in 1994. The small amount of overlap between the weighout and VTR series is illustrated in Figures C17 and C18. Splitting the weighout and VTR CPUE series can be justified by the differences in the way effort was measured and difference in the tilefish fleet between the series. In breaking up the series we omitted 1994 because there were very little CPUE data. The sparse 1994 data that existed came mostly from the weighout system in the first quarter of the year. Very similar trends exist in the four years of overlap between Turner (1986) CPUE and the weighout series (Figure C19).

A month vessel interaction was significant but explained only a small amount of the total sum of squares (6%*).* Adding a month - vessel interaction term to the GLM model had very little influence on the results (Figure C20). In addition, limiting the VTR series to the 5 dominant tilefish vessels also had little influence on GLM results. The GLM output for the weighout and VTR CPUE series standardized for individual vessel effects can be seen in Appendix C2 and C3.

Since 1979, the tilefish industry has changed from using cotton twine to steel cables for the backbone and from J hooks to circle hooks. In light of possible changes in catchability associated with these changes in fishing gear, the working group considered that it would be best to use the three available indices separately rather than combined into one or two series. The earliest series (Turner 1986) covered 1973-1982 when gear construction and configuration was thought to be relatively consistent. The Weightout series (1979-1993) overlapped the earlier series for four years and showed similar patterns (Figure C19) and is based primarily on catch rates from New Jersey vessels. The VTR (1995-2004) series is based primarily on information from New York vessels.

Commercial market category and size composition data

Six market categories exist in the database. From smallest to largest they are: small, kitten, medium, large and extra large as well as an unclassified category. In 1996 and 1997, the reporting of tilefish by market categories increased, with the proportion of unclassified catch declining to less than 20% (Table C10, Figure C21). The proportion of landings in the small and kitten market categories increased in 1995 and 1996. Small and kitten market categories had similar length distributions and samples were combined. Evidence of several strong recruitment events can be seen tracking through the market category proportions (Figures C21 and C22). The proportion of the large market category has declined since the early 1980s (Figure C22). Landings data obtained directly from the New York tilefish industry shows a similar decline in the proportion of the large market category between 1980 and 1990 (Figure C23).

Since 2000 commercial length samples from New York were measured in total length. All other commercial tilefish were measured in fork length. In 2005 port agents measured both total and fork length from 345 fish to determine a total to fork length conversion (Figure C24). A 45 cm fish has about a 2 cm difference between total and fork length. All total length measurement were converted to fork length using the total length to fork length regression.

Extensive size sampling was conducted in 1976-1982 (Grimes *et al.* 1980, Turner 1986) however that data are not available by market category. Since then commercial length sampling has been inadequate in most years (Table C10). However some commercial length sampling occurred in the mid to late 1990s. More recently there has been a substantial increase in the commercial length sampling in 2003 and 2004 (Table C10). Commercial length sampling in New York has also increased since the last assessment in 1998. The large and medium market category length frequencies appear to have been relatively stable for years when more than 100 fish were measured (Figures C25 and C26). However the small market category exhibits shifts in the size distribution in certain years as strong year classes move through the fishery (Figure C27). The tracking of a year class can be seen as the cohort grows over the year in 2002 and 2003 (Figure C28).

The loligo-scup small mesh trawl fishery catches smaller tilefish than longline gear. This can be seen in many of the length frequency distributions of smalls and kittens for the trawl gear (Figure C29). Therefore trawl length frequency distribution where not used to characterize the catch (Table C11). Longline tilefish fishermen often receive forecasts from the draggers of when a strong year class will be entering the fishery.

Commercial length frequencies were expanded for years where sufficient length data exist (1995-1999 and 2002-2004) (Table C10). The large length frequency samples from 1996 to 1998 were used to calculate the 1995 to 1999 expanded numbers at length while the large length samples from 2001 and 2003 were used to calculate the 2002 expanded numbers at length. Evidence of strong 1993 and 1999 year classes can be seen in the expanded numbers at length in the years when length data existed (1995-1999 and 2002-2004) (Figure C30). The matching of modes in the length frequency with ages was done using the Turner (1986) aging study. At the end of 2004 the 1999 year class can be seen growing into the medium market category (Figure C30). In recent years it appears that most of the catch is made up of this 1999 year class. An increase in the landings and CPUE can be seen when the 1993 and 1999 year classes recruit to the longline fishery.

Recently 1,409 commercial lengths were taken from 17 hauls on 3 tilefish longline observer trips from three different vessels (October 2004, November 2004, and January 2005) (Figure C31). The observer length frequency data show slightly larger fish than in the expanded commercial length data, which could be explained by growth of the cohort since the trips were done at the end of the year (Figure C32). A comparison between recent commercial expanded length data to commercial length data collected by Turner et al. (1983) from 1974-1982 shows a shift in the landings to smaller fish (Figure C33).

Recreational data

A small recreational fishery occurred briefly in the mid 1970s (< 100 mt annually, Turner 1986) but subsequent recreational catches have been quite low for the last 25 years (i.e., less than 1 mt caught annually) (Table C12). Party and charter boat vessel trip reports also show low numbers of tilefish being caught since 1994 (Table C13). Directed tilefish trips are rare. Since 2000, only 2 trips in the MRFSS data had tilefish reported as the primary target species.

NEFSC Trawl survey data

Only a few fish per survey are caught during NEFSC bottom trawl surveys. This survey time series is not useful as an index of abundance for tilefish.

TOR 2: ***Estimate fishing mortality and total stock biomass for the current year and characterize the uncertainty of those estimates.***

TOR 3: ***Evaluate and either update or re-estimate biological reference points as appropriate.***

3.0 MORTALITY AND STOCK SIZE ESTIMATES

Surplus production model

The ASPIC surplus production model (Prager 1994; 1995) was the primary model used to determine fishing mortality, stock biomass and biological reference points (F_{msy}, and B_{msy}). Results of sensitivity runs with 13 different configurations of the ASPIC model were examined (Table C14). A comparison of runs 1-2, 3-4, 5-6, and 7-8 provides information on the effect of splitting the weighout and VTR CPUE series. Runs 3-4, and 5-6 also extend the landings time series in the past before the existence of CPUE data. Runs 3-4 extended landings to the end of World War II (1945) when effort was thought to be low and runs 5-6 extended the landings to the beginning of the landings time series (1916). A comparison of runs 7-8 with runs 1-2 evaluates the effect of using a GLM to standardize CPUE. Runs 9 through 11 reduced the increase in CPUE at the end of the VTR series to determine the sensitivity of recent increases in CPUE to the model results (Figure C34). Run 12 examines the effect of using a single CPUE series by combining Turner and the weighout/VTR CPUE series. Turner and weighout-based CPUE indices were combined using a regression on the four years of overlap between the indices (1979-1982) (Figure C35). Run 13 fixed the $B1/B_{msy}$ ratio at 1.

Splitting of the weighout and VTR CPUE series did not have a strong effect on the model results. Extending the landings time series used in the model back to 1916 or 1945 when CPUE data do not exist also did not appear to influence the results. The use of a CPUE series standardized for vessels effects (GLM) produced little change in the results. Sensitivity runs that lowered the CPUE at the end of the VTR CPUE series had more of an influence on model results. Reducing the increase in CPUE at the end of the time series generally lowers the estimate of the intrinsic rate of increase. The sensitivity run that combined all of the CPUE series into a single index (run 12) provided a high estimate of the intrinsic rate of increase ($r = 0.63$). Large fluctuations in the $B1/B_{msy}$ ratio between the model runs did not have a large influence on model results. The Working Group accepted the formulation that began the analysis in 1973, separated the Turner, weighout and VTR CPUE into three series and fixed the $B1/B_{msy}$ ratio at 1 as the final run (run 13). The solution obtained from the final run was bootstrapped (1000 iterations) to obtain estimates of precision and bias. The complete ASPIC model output with bootstrap results is included as Appendix C4.

The surplus production model indicates that the tilefish stock biomass in 2005 has improved since the last assessment in 1998. Total biomass in 2005 is estimated to be

72% of B_{msy}, and fishing mortality in 2004 is estimated to be 87% of F_{msy} (Figure C36). Biological reference points did not change greatly from the 1998 assessment. B_{msy} is estimated to be 9,384 mt and F_{msy} is estimated to be 0.21 (Figure C37). Bootstrap iterations show highly variable estimates of 2005 total biomass to B_{msy} ratios (80% confidence intervals from 0.5 to 1.2) and 2004 F to F_{msy} ratios (80% confidence intervals from 0.5 to 1.3) (Figure C38, Appendix C4).

Catch-Length Model Mortality Estimates

A length-based fishing mortality estimate in the 1998 assessment for the 1996-1997 period was 0.65 using the Hoenig (1987) method and 1.12 using the Beverton and Holt (1957) method (Nitschke et al. 1998). In the present assessment a catch-length forward projection model was developed in an attempt to produce more accurate fishing mortality estimates based on growth and size information in the catch. Testing of the model produced reasonable results on a simulated population of tilefish when recruitment does not have a strong trend over time and the average growth is known. However the model could not fit both the catch length frequency and total landings data in the tilefish assessment. The model produced an unrealistic increase in F at the end of the time series. Substantial changes to model inputs (natural mortality, partial recruitment, and/or growth rate) were needed to eliminate the fitting conflict. The catch-length model was not considered as the primary model for determining stock status at this time because of the fitting problems and the uncertainty about the partial recruitment, natural mortality and growth. The expanded length frequency data for 2002-2004 indicates that most of the commercial landings were taken from a single year class (1999) comprising of relatively young fish (age 5 in 2004).

The longline tilefish fleet targets strong year classes by fishing areas where the catch rates are high. Spatial segregation of the stock by size and changes in fishing practices to keep catch rates high can result in a dome shaped partial recruitment pattern. The shape and changes over time of a possible dome is unknown. Assuming that natural mortality and growth are relatively well known, a severe dome shaped partial recruitment pattern is needed to allow fishing mortality to match the F trend seen in the ASPIC model. Conversely, if a flat top partial recruitment pattern is more likely to occur in the fishery, recent catches should have comprised more larger fish than were observed to allow the catch-length model to estimate a declining fishing mortality rate at the end of the time series. Although uncertainty in the input data and the paucity of length data from the fishery precluded the use of the catch-length model at this time, the model still calls attention to the lack of large fish seen in the catch in recent years for a stock which is thought to have a relatively low fishing mortality rate in recent years.

An Index Method (AIM)

An Index Method (AIM, NOAA Fisheries Toolbox V1.4.1) was used as an additional indicator of stock status. The Index Method can only accommodate a single CPUE series so the combined index was employed. AIM uses a statistical fitting procedure to determine the relationship between indices and landings to calculate a relative F. A replacement ratio is estimated by dividing the annual CPUE index by a moving average of the previous five years of that index. At a replacement ratio of 1 the stock is sustained at the same level as the previous five years. At a level above 1 the stock is increasing and at a level below 1 the stock is declining. A relative F is calculated by dividing the catch by the three-point moving average of the catch rates centered on the year in which that catch occurred. The relative F needed to maintain the population can be computed from the plot comparing the relative F with the replacement ratio (Figure C39).

For tilefish, the replacement ratio has been increasing since 2001 and has been above 1.0 since 2002, and the current estimate of relative F for 2004 is well below the point corresponding to the replacement ratio of 1.0 (Figure C40, Appendix C5). This model indicates that relative F has declined in recent years (Figure C40).

Lagged Recruitment Survival Growth (LRSG) Model

A lagged-recruitment survival growth (LRSG) model (Hilborn and Mangel 1997) was developed for tilefish. This simple model includes a time lag for recruitment (L) and a lumped survival-growth parameter for biomass (s). The model was fit using catch biomass and combined catch-per-unit effort (CPUE) series during 1973-2004. The recruitment time lag was 4 years. Recruited biomass in year T+1 (B_{T+1}, age-4+) was derived from previous biomass, recruiting biomass (R_T), and catch (C_T) via

$$B_{T+1} = s \cdot B_T + R_T - C_T$$

Recruitment biomass was modeled using a Beverton-Holt curve with a time lag of L=4 years

$$R_T = \frac{B_{T-L}}{a + b \cdot B_{T-L}}$$

In the likelihood for CPUE, model observation errors were assumed to be iid (independent and identically distributed) multiplicative lognormal distributions with constant variance. CPUE was assumed proportional to age-4+ biomass raised to an exponent (δ). In practice, there was insufficient information to estimate δ and it was set

$$CPUE_T = q \cdot \left(B_T\right)^{\delta}$$

to unity.
Prior distributions were assumed to be uninformative, with the exception of stock-recruitment steepness. Broad uniform prior distributions were used for the initial biomass (B_0), survival (s), catchability (q), exponent (δ), and error variance (σ^2) parameters. A uniform prior of [0.2, 1] was initially used for the stock-recruitment steepness parameter (z). This initial model configuration led to a highest posterior density point estimate of z=0.88 indicating a highly resilient stock. However, the Hessian matrix for this model solution had a high condition number indicating substantial collinearity among

parameters. As a result, an informative truncated Gaussian prior for steepness was developed using the meta-analysis of Myers et al. (1999). Steepness estimates from the nearest taxonomic grouping were used to set the mean steepness for the prior. In this case, the closest group was striped bass (*Morone saxatalis*) with a steepness of z=0.82. The coefficient of variation for the steepness prior was assumed to be 20%. Realized steepness values constrained to be in the interval [0.2, 1.0].

The combined CPUE series was used, because the current configuration of the model allows only one index of abundance. The LRSG model provided a reasonable fit to the CPUE series (Figure C41). Standardized residuals (Figure C42) were smaller than 1.5 and they exhibited a moderate alternating high-low pattern across blocks of several years. Relative biomass estimates (B/B_{msy}) indicated that the tilefish stock had been fished down in the 1970s-1980s (Figure C43) and has moderately increased since then. Recent biomass estimates appear to be at or above the B_{msy} estimate obtained from this model. Relative exploitation rate estimates (H/H_{msy}) indicated that the tilefish stock experienced periods of overfishing during the 1980s-1990s (Figure C44). Recent exploitation rates appear to be relatively low but increasing. Overall the LRSG modeling results are more similar to the results obtained from the ASPIC model calibrated with the single linked CPUE series.

Yield and Spawning Stock Biomass per Recruit

Biological reference points from the Thompson-Bell yield per recruit (YPR) model (Thompson and Bell 1934) were not updated from the last assessment since updated data for the YPR analysis does not exist. However a value of F_{max} was calculated from the Catch-length model. A length based YPR analysis (NOAA Fisheries Toolbox V1.2.1) was also performed for comparison to F_{max} estimates derived from the Catch-length model and the original 1998 YPR analysis. The proportions mature-at-age and length were derived from estimates of maturity in 1978 and 1982 provided by Grimes et al. (1988) (Figure C45). In the 1998 YPR analysis the partial recruitment and weight at age was taken from the yield per recruit analysis (Ricker model) in Turner (1986). Von Bertalanffy growth parameters, a length weight relationship and a partial recruitment vector based on the landings length frequencies are used in the catch-length model and length based YPR model. The 1998 yield per recruit analysis provided an estimate of F_{max} = 0.143, the length based YPR model provided an estimate of 0.138 (Figure C46, Appendix C6) and the catch-length model estimated an F_{max} of 0.142 (Figure C47). The predicted length and age distribution at F_{max} from the catch-length model is shown in Figure C48.

TOR 4: ***Where appropriate, estimate a constant TAC and/or TAL based on stock status for years following the terminal assessment year.***

TOR 5: ***If projections are possible,***

a) provide seven year projections of stock status under various TAC strategies and

b) evaluate current and projected stock status against existing rebuilding or recovery schedules, as appropriate.

4.0 Biomass and Fishing Mortality Projections

The Working Group examined several ASPIC projections employing a constant TAC strategy, including the current TAC of 905 mt. Each of these analyses exhibited wide variance and substantial bias and, in many cases, produced estimates of biomass and F at maximum or minimum model boundary conditions. The Working Group, therefore, concluded that the projections are too uncertain to form the basis for evaluating likely biomass recovery schedules relative to Bmsy under various TAC strategies. We do note, however, that stock biomass in 2005 (72% of B_{msy}) is above that projected for 2005 in the 1998 assessment (59% of B_{msy}). Thus, the existing TAC of 905 mt appears to have sufficiently constrained F to allow stock biomass to increase towards B_{msy}.

5.0 CONCLUSIONS

The Working Group accepted the ASPIC model solution but the projection results were considered too uncertain to form the basis for evaluating likely biomass recovery schedules relative to B_{msy} under various TAC strategies. The surplus production model indicates that the tilefish stock biomass in 2005 has improved since the last assessment in 1998. Total biomass in 2005 was estimated to be 72% of B_{msy} and fishing mortality in 2004 was estimated to be 87% of F_{msy}. MSY and Yield per recruit based biological reference points did not change greatly from the 1998 assessment. Results from the AIM model suggest that relative F is below the point that corresponds with a replacement ratio of 1.0 (stock replacement) and the LRSG model produced results similar to the ASPIC surplus production model. The AIM and LRSG require a single index of abundance. The ASPIC model, which allows for the separation of the CPUE indices, was used as the base model for status determination given the changes in commercial gear over time. However commercial length data indicate that improvements in total biomass are predominantly due to a strong 1999 year class. Most of the commercial catch was derived from this year class over the 2002-2004 period.

The partial recruitment pattern is unknown for the tilefish longline fishery because targeting of year classes to increase catch rates and market conditions will influence the size of fish landed. The price on the large market category in this fishery is particularly sensitive to the quantity of large fish landed. However there is still concern that fishing mortality may be higher than estimated by the surplus production model due to the relative lack of larger/older fish seen in the catch. The inability to characterize the actual partial recruitment pattern, the possibility of unknown refuge effects due to conflicts with lobster and trawl gear and effects of targeting incoming year classes introduce considerable uncertainty in interpreting CPUE from this fishery as a measure of stock abundance. Thus, there is concern that CPUE at the end of the series may be increasing faster than stock biomass. CPUE and catch length frequency data in this fishery may be as much a reflection of changes in fishing practices and the spatial distribution of the fish rather than fluctuations in population size.

With regard to the yield per recruit-based reference points and the results from the catch-length model, there is an issue of how appropriate it is to assume a flat top partial recruitment pattern given anecdotal information that the tilefish fleet will target single year classes and will optimize profits by fishing an area where the catch rates are higher on fish in the small and medium market category as opposed to an area (greater depth) where more valuable larger fish can be caught at a lower catch rate.

6.0 SOURCES OF UNCERTAINTY

There are two major sources of uncertainty affecting our perception of current stock status. The biomass-based models (ASPIC, AIM and LRSG) use the CPUE series as an index of population size. The Working Group considered these models and expressed concerns over whether the CPUE in this fishery may be as much a reflection of changes in fishing practices and changes in spatial distribution of the fish rather than fluctuations in population size. The catch-length model attempts to reconcile recent fishing mortality rates with a less than expected representation of larger fish in the catch. Because there are no fishery-independent data on trends in population biomass and size structure, the model must assume that the length composition of the catch will represent the extent of large fish in the population assuming a flat topped partial recruitment pattern. Specific sources of uncertainty are:

1) The effort metric (days absent) in the Weighout and VTR CPUE is a crude measure of effort and could be improved by collecting information (number and size of hooks, length of main line, soak time, time of day, depth fished and area fished) on a haul by haul basis and not by a trip basis.

2) The production models and index method (AIM) do not consider size or age structure of the population.

3) Sparse commercial length frequency sampling in many years.

4) The possible existence of a dome shaped partial recruitment pattern in the longline fishery depending on hook size and/or fishery practice such as areas/depth fished.

5) Possible shifts in growth relative to the Turner (1986) study and maturity at age/size from the Grimes *et al.* (1988) early 1980s study with increases in fishing mortality in the 1990s.

6) Effects of fishing on spawning success for a species that possesses sexual dimorphic growth and size specific competition for baited hooks.

7) Effects of fish behavior and fishing practice on the CPUE index as an assumed measure of population size.

8) Uncertainty in projections based on wide variance and substantial bias estimates.

7.0 RESEARCH RECOMMENDATIONS

1) Conduct a hook selectivity study to determine partial recruitment changes with hook size. Determine catch rates by hook size. Update data on growth, maturity, size structure, and sex ratios at length.

2) Collect data on spatial distribution and population size structure. This can help answer the question of the existence of a possible dome shaped partial recruitment pattern where larger fish are less vulnerable to the fishery due to spatial segregation by size.

3) Continue to develop the forward projecting catch-length model as additional length data becomes available. Investigate the influence of adding a tuning index of abundance and model estimated partial recruitment (logistic) to the catch-length model.

4) Collect appropriate effort metrics (number and size of hooks, length of main line, soak time, time of day, area fished) on a haul basis to estimate commercial CPUE.

5) Initiate a study to examine the effects of density dependence on life history parameters between the 1978-82 period and present.

6) Increased observer coverage in the tilefish fishery to obtain additional length data.

7) Develop a bioeconomic model to calculate maximum economic yield per recruit.

TOR 6: Review, evaluate and report on the status of the research recommendations offered in the 1999 Science and Statistical committee reviewed assessment.

Research recommendations from 1999 Science and Statistical Committee review

1) Ensure that market category distributions accurately reflect the landings.

This is not really a research recommendation. The catch-length model assumes that landings from all market categories are accurately accounted for and that the length frequency distributions for a market category are stable over time. Sampling of the commercial lengths has improved over the last two years.

2) Ensure that length frequency sampling is proportional to landings by market category.

This is not really a research recommendation. Commercial length sampling has been sporadic over the time series. In particular length samples from the large market category have been lacking. However commercial length sampling improved in 2003 and 2004.

3) Increase and ensure adequate length sampling coverage of the fishery.

Commercial length sampling improved in 2003 and 2004.

4) Update age- and length- weight relationships.

This TOR has not been addressed. Question why length-weight relationships would change. Growth data for tilefish should be updated and will be collected in a planned 2005-2006 hook selectivity study.

5) Update the maturity-at-age, weight-at-age, and partial recruitment patterns.

This TOR has not been addressed. Maturity and partial recruitment data will also be collected in the 2005-2006 hook selectivity study.

6) Develop fork length to total length conversion factors for the estimation of total length to weight relationships.

This work is in progress. Port agents are collecting data.

7) Incorporate auxiliary data to estimate r independent of the ASPIC model.

This TOR has not been addressed. Question if this can be done or should be done.

8.0 REFERENCES

Beverton, R.J.H. and S.J. Holt. 1957. On the dynamics of exploited fish populations. Fish. Invest. Minist. Agric. Fish. Food G.B. (2. Sea Fish.) 19:553.

Freeman, B.L., and S.C. Turner. 1977. Biological and fisheries data on tilefish, *Lopholatilus chamaeleonticeps* Goode and Bean. NEFSC Tech. Rep. 5.

Granger, C.W. and P. Newbold. 1977. Forecasting Economic Time Series. New York: Academic Press.

Grimes, C.B., K.W. Able, and S.C. Turner. 1980. A preliminary analysis of the tilefish, *Lopholatilus chamaeleonticeps*, fishery in the Mid-Atlantic Bight. Mar. Fish. Rev. 42(11): 13-18.

Grimes, C.B., C.F. Idelberger, K.W. Able, and S.C. Turner. 1988. The reproductive biology of tilefish, *Lopholatilus chamaeleonticeps*, Goode and Bean, from the United States Mid-Atlantic Bight, and the effects of fishing on the breeding system. Fish. Bull. 86(4): 745-762.

Hilborn, R., and M. Mangel. 1997. The ecological detective. Princeton University Press, Princeton, NJ, 315 pp.

Hoenig, J.M. 1987. Estimation of growth and mortality parameters for use in length-structured stock production models. Pages 121-128 *in* Length-based methods in fisheries research, edited by D. Pauley and G.R. Morgan. Int. Center for Living Aquatic Resource Management. ICLARM Conference Proceedings 13, 468.

Mayo, R.K., T.E. Helser, L. O'Brien, K.A. Sosebee, B.F. Figuerido, and D. Hayes. 1994. Estimation of standardized otter trawl effort, landings per unit effort, and landings at age for Gulf of Maine and Georges Bank cod. NEFSC Reference Document 94-12.

Mid-Atlantic Fishery Management Council. 2000. Tilefish fishery management plan. NOAA award No. NA57FC0002 .

Myers, R.A., Bowen, K.G., and N.J. Barrowman. 1999. Maximum reproductive rate of fish at low population sizes. Canadian Journal of Fisheries and Aquatic Sciences 56: 2404-2419.

Northeast Fisheries Science Center (NEFSC). 1993. Report of the 16th Northeast Regional Stock Assessment Workshop (16th SAW) Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NEFSC Reference Document 93-18. 108 p.

Nitschke, P., G. Shepherd, and M. Terceiro. 1998. Assessment of tilefish in the middle Atlantic – southern New England region. NEFSC. 1-12.

Prager, M.H. 1994. A suite of extensions to a nonequilibrium surplus-production model. Fish. Bull. 92: 374-389

Prager, M.H. 1995. Users manual for ASPIC: a stock-production model incorporating covariates. SEFSC Miami Lab Doc. MIA-92/93-55.

Thompson. W.F. and F. H. Bell. 1934. Biological statistics of the Pacific halibut Fishery. 2. Effect of changes in intensity upon total yield and yield per unit of gear. Report of the International Pacific Halibut Commission 8: 49 p.

Turner, S.C., C.B. Grimes, and K.W. Able. 1983. Growth, mortality, and age/size structure of the fisheries for tilefish, *Lopholatilus chamaeleonticeps*, in the Middle Atlantic-Southern New England region. Fish. Bull. 81(4): 751-763.

Turner, S.C. 1986. Population dynamics of and, impact of fishing on tilefish, *Lopholatilus chamaeleonticeps*, in the Middle Atlantic-Southern New England region during the 1970's and early 1980's. New Brunswick, N.J.: Rutgers University. Ph.D. dissertation.